

[illegible]

CATCTCTTGT	TCTCTCCGCC	CATCTCTGCT	CTCTTTTATT	TTCCCAGAAA	GTTTTTTTTT	60
TTTTTTCCGA	ATTCGGTTAA	TCTCATTGGG	GTTTCCATTG	ATAGCAATGG	CGACGGCTTT	120
CGCTCCCACT	AAGCTCACTG	CCACGGTTCC	TCTGCATGGA	TCCCATGAGA	ATCGTCTCTT	180
GCTCCCGATC	CGATTGGCTC	CTCCTTCTTC	TTTCCTCGGA	TCCACCCGTT	CCCTCTCCCT	240
TCGCAGACTC	AATCACTCCA	ACGCCACCCG	TCGATCTCCC	GTCGTCTCTG	TCCAGGAAGT	300
TGTCAAGGAG	AAGCAATCCA	CCATAAATAC	CAGCCTGTTG	ATAACCAAG	AGGAAGGATT	360
GGAGTTGTAT	GAAGATATGA	TACTAGGTAG	ATCTTTCGAA	GACATGTGTG	CTCAAATGTA	420
TTACCGGAGC	AAGATGTTTG	GTTTTGTTCA	CTTGTA CAAT	GGCCAAAGAG	CTGTTTCTAC	480
TGGCTTTATC	AAGCTCCTTA	CCAAGTCTGA	CTCTGTCGTT	AGTACCTACC	GTGACCATGT	540
CCATGCCCTC	AGCAAAAGTG	TCTCTGCTCG	TGCTGTTATG	AGCGAGCTCT	TCGGCAAGGT	600
TACTGGATGC	TGCAGAGGCC	AAGGTGGATC	CATGCACATG	TTCTCCAAAG	AACACAACAT	660
GCTTGGTGGC	TTTGCTTTTA	TTGGTGAAGG	CATTCTCTGC	GCCACTGGTG	CTGCCCTTAG	720
CTCCAAGTAC	AGGAGGGAAG	TCTTGAAACA	GGATTGTGAT	GADGTCACTG	TCGCCCTTTT	780
CGGAGATGGA	ACTTGTAACA	ACGGACAGTT	CTTCGAGTGT	CTCAACATGG	CTGCTCTCTA	840
TAAACTGCCCT	ACTAGCTTTG	TTGTGAGAG	TAACCTTGTG	GCCATTGGGA	TGTCTCACTT	900
GAGAGCCACT	TCTGACCCCG	AGATTTGGAA	GAAAGGTCTT	GCATTTGGGA	TGCCCTGGTG	960
TCATGTTGAC	GGTATGGATG	TCTTGAAGGT	CAGGGAAGTC	GCTAAAGAAG	CTGTCACTAG	1020
AGCTAGAAGA	GGAGAAGGTC	CAACCTTGTT	TGAATGTGAG	ACTTATAGAT	TCAGAGGACA	1080
CTCCTTGGCT	GATCCCGATC	AGCTCCGTGA	TGCTGCTGAG	AAAGCCAAAT	ACGCGGOTAG	1140
AGACCCAAAT	GCAGCATTGA	AGAAGTATTT	GATAGAGAAC	AAGCTTGCAA	AGGAAGCGAG	1200
GCTAAAGTCA	ATAGAGAAAA	AGATAGACGA	GTTGGTGGAG	GAAGCGGTTG	AGTTTGCAGA	1260
CGCTAGTCCA	CAGCCCGGTC	GCAGTCAGTT	GCTAGAGAAAT	GTTGTTGCTG	ATCCAAAAGG	1320
ATTTTGAATT	GGACCTGATG	GACGGTACAG	ATGTGAGGAC	CCCAAGTTTA	CCGAAGGCAC	1380
AGCTCAAGTC	TGAGAGAGACA	AGTTTAACCA	TAAAGCTGCT	ACTGTCTCTT	CGATGTTTCT	1440
ATATATCTTA	TTAAGTTAAA	TGCTACAGAG	AATCAGTTTG	AATCATTTCG	ACTTTTGTCT	1500
TAAAAA	AAAAA	AAAAA				1530

# APPENDIX B

<u>MATAFAPTKL TATVPLHGSH ENRLLLPIRL APPSSFLGST RSLSLRRLNH SNATRRSPVV</u>	60
<u>SVQEVVKEKQ</u> STNN†SLLIT KEEGLELYED MILGRSFEDM CAQMYRGKM FGFVHLYNGQ	120
EAVSTGFIKL LTKSDSVVST YRDHVHALSK GVSARAVMSE LFGKVTGCCR GQGGSMHMF	180
KEHNMLGGFA FIGEGIPVAT GAAFSSKYRR EVLKQDCDDV TVAFFGDGTC NNGQFFECLN	240
MAALYKLPPII FVVENNLWAI GMSHLRATSD PEIWKKGPAF GMPGVHVDGM DVLKVREVA	300
EAVTRARRGE GPTLVECETY RFRGHSLADP DELRDAAEKA KYAARDPIAA LKKYLIENKL	360
AKEAELKSIE KKIDELVEEA VEFADASPQP GRSQLENVF ADPKGFGIGP DGRYRCEDPK	420
EGTAOV	428

# APPENDIX C

GAAAAAATGT	CTTCGATAAT	CCATGGAGCT	GGAGCTGCTA	CGACGACGTT	ATCGACGTTT	60
AATTCGTCG	ATTCCAAGAA	ACTCTTCGTT	GCTCCTTCTC	GCACAAATCT	TTCAGTGAGG	120
AGCCAGAGAT	ATATAGTGGC	TGGATCTGAT	GCGAGTAAGA	AGAGCTTTGG	TTCTGGACTT	180
AGAGTTCGTC	ACTCTCAGAA	ATTGATTCCA	AATGCTGTTG	CGACGAAGGA	GGCGGATACG	240
TCTGCGAGCA	CTGGACATGA	ACTATTGCTT	TTCGAGGCTC	TTCAGGAAGG	TCTGGAAGAA	300
GAGATGGACA	GAGATCCACA	TGTATGTGTT	ATGGGTGAAG	ATGTTGGCCA	TTACGGAGGT	360
TCCTACAAGG	TAACCAAGG	CCTTGCTGAT	AAATTTGGTG	ACCTCAGGGT	TCTCGACACT	420
CCTATTTGTG	AAAATGCATT	CACCGGTATG	GGCATTGGAG	CTGCCATGAC	TGGTCTAAGA	480
CCCGTTATTG	AAGGTATGAA	CATGGGTTTC	CTCCTCCTCG	CCTTCAACCA	AATCTCCAAC	540
AACTGTGGAA	TGCTTCACTA	CACATCCGGT	GGTCAGTTTA	CGATCCCGGT	TGTCATCCGT	600
GGACCTGGTG	GAGTGGGACG	CCAGCTTGGT	GCTGAGCATT	CACAGAGGTT	AGAATCTTAC	660
TTTCAGTCCA	TCCCTGGGAT	CCAGATGGTT	GCTTGCTCAA	CTCCTTACAA	CGCCAAAGGG	720
TTGATGAAAG	CCGCAATAAG	AAGCGAGAAC	CCTGTGATTC	TGTTTGAACA	CGTGCTGCTT	780
TACAATCTCA	AGGAGAAAAT	CCCGGATGAA	GATTACATCT	GTAACCTTGA	AGAAGCTGAG	840
ATGGTCAGAC	CTGGCGAGCA	CATTACCATC	CTCACTTACT	CGCGAATGAG	GTACCATGTG	900
ATGCAGGCAG	CAAAAACCTCT	GGTGAACAAA	GGGTATGACC	CCGAGGTTAT	CGACATCAGG	960
TCACTGAAAC	CGTTCGACCT	TCACACAATT	GGAAACTCGG	TGAAGAAAAC	ACATCGGGTT	1020
TTGATCGTGG	AGGAGTGTAT	GAGAACCGGT	GGGATTGGGG	CAAGTCTTAC	AGCTGCCATC	1080
AACGAGAACT	TTCATGACTA	CTTAGATGCT	CCGGTGATGT	GTTTATCTTC	TCAAGACGTT	1140
CCTACACCTT	ACGCTGGTAC	ACTGGAGGAG	TGGACCGTGG	TTCAACCGGC	TCAGATCGTG	1200
ACCGCTGTCT	AGCAGCTTTG	CCAGTAAATT	CATATTTATC	CGATGAACCA	TTATTTATCA	1260
TTTACCTCTC	CATTTCTTTT	CTCTGTAGCT	TAGTTCTTAA	AGAATTTGTC	TAAGATGGTT	1320
TGTTTTTGTG	AAAGTTTGTG	TCCTTTGTGG	TGTCTTTTAA	TATGGTTTGT	AACTCAGAAAT	1380
GTTTGTGTTG	TAATTTTATC	TCCCACTTTC	TTTTAAAAAA	AAAAAATAAA	AAAAAATAAA	1440
						1441

[illegible]

# APPENDIX E

<u>GGCGGATCTG GTTTCCTAGA TCCAAAACCC TTGTTTCTAG CTTCAGACAT</u>	50
<u>AATCTAAATT TGTCGACAAT TCTCATAAAA CGTGATTACT CTCATCGTCC</u>	100
<u>CATCTTCTAT ACAACTTCTC AGTTATCTTC AACGGCGTAT TTGAGTCCCT</u>	150
<u>TCCGTAGCCT CCGTATGAG TCTACGCCCC TGGAGACACA GGCTGATCAT</u>	200
<u>TTGGTTCAGC AGATTGATGA AGTCGATGCC CAGGAAGTGG ATTTCCCAGG</u>	250
<u>AGGCAAAGTC GGTACACAT CGGAGATGAA ATTACATACCG GAATCATCTT</u>	300
<u>CAAGCAGGAT TCCATGTTAC CCGGTCTTG ACCAAGACGG ACCAATCATC</u>	350
<u>CCCGATAGCG ATTTTATTCC GGTGAGTGAG AAATCGGTG TTAGAATGTA</u>	400
<u>CGAACAAATG CCGACGCTAC AAGTAATGGA TCACATCTTC TACGAAGCTC</u>	450
<u>AACGTCAAGG AAGAATATCT TTTTATCTTA CTTCGGTCCG ACAAGAAGCC</u>	500
<u>ATTAACATCG CTTCAGCACC TGCTCTCAGT CCTGACGACG TCGTTTACC</u>	550
<u>TCAGTACCGA GAACCTGGAG TTCTTTTGTG GCGTGGCTTC ACGTTGAGG</u>	600
<u>AGTTTCTAA TCAGTGTTTT GGGACAAAG CTGATTATCG CAAAGGCAGA</u>	650
<u>CAAATGCCAA TTCAATTACCG TTCCAATCGT CTTAATTACT TCACTATCTC</u>	700
<u>GTCTCCAATT GCCACGCAAC TTCTCAAGC TGCTGGAGTT GGTATTCTT</u>	750
<u>TGAAAATGGA CAACAACAAT GCTTGTACTG TTACATTCAT CCGACATCGT</u>	800
<u>GGCACAAGCG AGCGAGATTT TCACGCGCGA TTGAATTTTG CGCCCGTAAT</u>	850
<u>GGAACCTCCG GTTGTGTTTA TATGTGCGAA CAACGGTTCG GCGATTAGTA</u>	900
<u>CTCATATCTC AGAACAGTTT ACAAGTCATG CAATAGTTGT GAAAGGTCAA</u>	950
<u>GCTTACGGTA TCCCGAAGCA TCCCGTGTGG CACGGTACCG ATGCACTTGC</u>	1000
<u>GGTTTATAGT GCTGTACGCT CAGCTCGAGA AATGGCTGTA ACAGAACAAA</u>	1050
<u>GACCTGTCT CATTCAGATG ATGACATATA GAGTAGGACA TCATTCTACA</u>	1100
<u>TCAGATGATT CAACTAAGTA CAGGGCGCGG GATCAAATCC AGTACTGGAA</u>	1150
<u>AATGTCCAGA AACCTGTGA ATACATTTCC GAAATGGGTC GAAGATAACG</u>	1200
<u>CAITGGTGGAG TCAGGAAGAT GAATCCAAGC TAACATCTAA CGCAACAAAA</u>	1250
<u>CAGCTTCTGC AACCGATTCA GGCTGCGGAG AAGTGGGAGA AACCAACCATT</u>	1300
<u>CACAGAGTTG TTTAACCATG TATATGATGT TAAACCGAAG AACCTAGAAG</u>	1350
<u>AGCAACAAC TGGTTTGAAG GAATTAGTAA AGAACAACC TCAACATTAT</u>	1400
<u>CCTCCTGGCT TTCAATGTTG AATCTAGAGG AACTGTGTGG TTAAAATACC</u>	1450
<u>TCCCGGACCG CGAATTCGAT ATCAACCTTC TCATTGCAGA CTATTTATAT</u>	1500
<u>TGTCCACGTA TCGAATAGTA ATCAAGTATC AATGTAGAGA CCAGCATTTG</u>	1550
<u>GAGCATCAAA AAAAAAAAAA AAAAAAAAAA AAAAAA</u>	1587

# APPENDIX F

AIWFARSKTL VSSLRHNLNL STILIKRDYS HRPIFYTTSD LSSTAYLSPF	50
GSLRHESTAV ETQADHLVQQ IDEVDAQELD FPGGKVGYS EMKFIPSSS	100
RRIPCYRVLD EDGRIIPDSO FIPVSEKLAV RMYEQMATLO VMDHIFYEAO	150
QROGRISFYLT SVGEEAINIA SAAALSPDDV VLPQYREPGV LLWRGFTLEE	200
TPP binding site	
QANQCFGNKA DYGKGROMPI HYGSNRLNYF TISSPIATQL PQAAGVGYSL	250
BCOADC E18 binding site	
QMDKKNACTV TFIGDGGTSE GDFHAGLNFA AVMEAPVVFI CRNNGWAIST	300
QHISEQFRSDG IVVKGQAYGI PKHPVWDGTD ALAVYSAVRS AREMAVTEQR	350
QVLIEMMTYR VGGHSTSDS TKYRAADEIQ YWKMSRNPVN RFRKWVEDNG	400
QWSEEDESKL RSNARKOLLO AIQAAEKWEK QPLTELFNDV YOVKPKNLEE	450
QELGLKELVK KQPDYPPGF HV	472

# APPENDIX G

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
TTCCTCACCC	ACCAAAAGTA	GCAAACCTTT	GCCACCTAAA	AATCTTACCA	50
GTTGGGIGAA	AGTTGCCAAA	ATAGAGCTTG	CTTTTGTCGC	AATCCTATAT	100
TTTTCAGATT	GATTGTTGGT	GGGTTTGIGT	AAATGGCGGC	TCTTTTAGGC	150
ACATCCTGCC	GGAAACTGAG	TTTTCCGAGC	TTGACTCAGC	GAGCTAGGAG	200
GGTATCGACG	GAAACTGGAA	AACCATTGAA	TCTATACTCT	GCTATTAATC	250
AAGCGCTTCA	CATCGCTTTG	GACACCGATC	CTCGGTCCTA	TGTCCTTGGG	300
GAACACGTTG	GCTTTGGTGG	AGTCTTTCGC	TGTACAACCTG	GTTTAGCTGA	350
ACGATTCGGG	AAAAACCGTG	TCTTCAATAC	TCCCTCTTGT	GAGCAGCGCA	400
TGTGTGGATT	TGGCATTGGT	CTAGCAGCAA	TGGGTAATCG	AGCAATTGTA	450
GAGATTCAGT	TTGCAGATT	TATATATCCT	GCTTTTGATC	AGATTGTTAA	500
TGAAGCTGCA	AAGTTCAGAT	ACCGAAGTGG	TAACCAATTC	AACGTGTGGAG	550
GACTTACCAT	AAGAGCACCA	TATGGAGCAG	TTGGTCATGG	TGGACATTAC	600
GATTCACAAT	CCCCTGAAGC	TTTCTTTTGC	CATGTCCCTG	GTATTAAGGT	650
TGTTATCCCT	CGGAGTCCAC	GAGAAGCAAA	GGGACTGTTG	TTGTTCATGTA	700
TCCGTGATCC	AAATCCCGTT	GTTTCTTTCG	AACCAAAGTG	GCTGTATCGT	750
CAAGCAGTAG	AAGAAGTCCC	TGAGCATGAC	TATATGATAC	CTTTATCAGA	800
AGCAGAGGTT	ATAAGAGAAG	GCAATGACAT	TACACTGGTT	GCATGGGGAG	850
CTCAGCTTAC	CGTTATGCAA	CAAGCTTGTC	TGGACGCGGA	AAAGGAAGGA	900
ATATCATGTG	AACTGATAGA	TCTCAAGACA	CTGCTTCCTT	GGCACAAGA	950
AACCGTTGAG	GCTTCAGTTA	AAAAGACTGG	CAGACTTCTT	ATAAGCCATG	1000
AAGCTCCTGT	AACAGGAGGT	TTTGGAGCAG	AGATCTCTGC	AACAATTCTG	1050
GAACGTTGCT	TTTTGAAGTT	AGAAGCTCCA	GTAAGCAGAG	TTTGTGGTCT	1100
GGATACTCCA	TTTCCTCTTG	TGTTTGAACC	ATTCTACATG	CCCACCAAGA	1150
ACAAGATATT	GGATGCAATC	AAATCGACTG	TGAATTACTA	GCCGTACTAT	1200
CTGTAGTTTA	CTGTTTACAC	TAGCACTAAT	GTAATCCCAT	GTCTTTGTTA	1250
TCAATTCGTC	TAATGTAACA	CTACCGATT	ACTTTAATGA	ATTTCAAGAT	1300
AACCAAAAAA	AAAAAAA				1319

1967		1968		1969		1970		1971		1972		1973		1974		1975		1976		1977		1978		1979		1980		1981		1982		1983		1984		1985		1986		1987		1988		1989		1990		1991		1992		1993		1994		1995		1996		1997		1998		1999		2000		2001		2002		2003		2004		2005		2006		2007		2008		2009		2010		2011		2012		2013		2014		2015		2016		2017		2018		2019		2020		2021		2022		2023		2024		2025		2026		2027		2028		2029		2030		2031		2032		2033		2034		2035		2036		2037		2038		2039		2040		2041		2042		2043		2044		2045		2046		2047		2048		2049		2050		2051		2052		2053		2054		2055		2056		2057		2058		2059		2060		2061		2062		2063		2064		2065		2066		2067		2068		2069		2070		2071		2072		2073		2074		2075		2076		2077		2078		2079		2080		2081		2082		2083		2084		2085		2086		2087		2088		2089		2090		2091		2092		2093		2094		2095		2096		2097		2098		2099		2100		2101		2102		2103		2104		2105		2106		2107		2108		2109		2110		2111		2112		2113		2114		2115		2116		2117		2118		2119		2120		2121		2122		2123		2124		2125		2126		2127		2128		2129		2130		2131		2132		2133		2134		2135		2136		2137		2138		2139		2140		2141		2142		2143		2144		2145		2146		2147		2148		2149		2150		2151		2152		2153		2154		2155		2156		2157		2158		2159		2160		2161		2162		2163		2164		2165		2166		2167		2168		2169		2170		2171		2172		2173		2174		2175		2176		2177		2178		2179		2180		2181		2182		2183		2184		2185		2186		2187		2188		2189		2190		2191		2192		2193		2194		2195		2196		2197		2198		2199		2200		2201		2202		2203		2204		2205		2206		2207		2208		2209		2210		2211		2212		2213		2214		2215		2216		2217		2218		2219		2220		2221		2222		2223		2224		2225		2226		2227		2228		2229		2230		2231		2232		2233		2234		2235		2236		2237		2238		2239		2240		2241		2242		2243		2244		2245		2246		2247		2248		2249		2250		2251		2252		2253		2254		2255		2256		2257		2258		2259		2260		2261		2262		2263		2264		2265		2266		2267		2268		2269		2270		2271		2272		2273		2274		2275		2276		2277		2278		2279		2280		2281
------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
MAALLGRSCR	KLSFPSLTHG	ARRVSTETGK	PLNLYSAINQ	ALHIALDTDP	50
RSYVFGEDVG	FGGVFRCTTG	LAERFGKNRV	FNTPLCEQGI	VGFGIGLAAM	100
GNRAIVEIQF	ADYITPAFDQ	IVNEAAKFRY	RSGNQFNCGG	LTIAPYCAV	150
GHGGHYHSQS	PEAFFCHVPG	IKVVI PRSPR	EAKGLLLSCI	RDPNFVVFEE	200
PKWLYRQAVE	EVPEHDYMIP	LSEAEVIREG	NDITLVGWA	QLTVMQACL	250
DAEKEGISCE	LIDLKTL LPW	DKETVEASVK	KTGRLLISHE	AFVTGGFGAE	300
ISATILERCF	LKLEAFVSRV	CGLDTPFPLV	FEPFYMPKN	KILDAIKSTV	350
NY					352